

3. 2

09-219835

(43)Date of publication of application : 19.08.1997

H04N 5/765
H04N 5/781
G06F 17/30
H04N 5/91

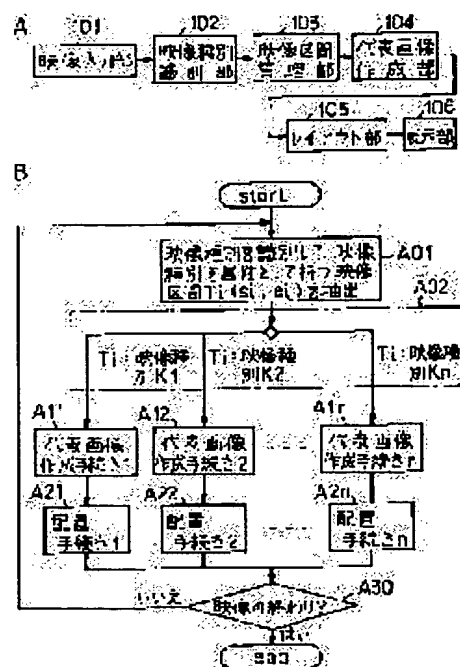
(71)Applicant : NIPPON TELEGR & TELEPH CORP
<NTT>

(72)Inventor : TANIGUCHI YUKINOBU
HAMADA HIROSHI

(57)Abstract:

PROBLEM TO BE SOLVED: To extract a typical picture in a method proper for the each class of video.

SOLUTION: A video class identification part 102 identifies the video class of video data inputted from a video inputting part 101 to extract a video section T_i (s_i and e_i) with the video class as an attribute. A video section managing part 103 manages the extracted video section. A typical video preparing part 104 executes a procedure previously given for each video section to extract or compose a picture representing each video section T_i . A layout part 105 prepares a video summary arranging the typical pictures spatially and a display part 16 displays it. It is possible to select commercial (CM) as one of the video classes to extract a final picture as the typical picture with respect to the video section (CM clip) with CM as its attribute.



[Date of request for examination]	28.10.1999
[Date of sending the examiner's decision of rejection]	
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]	
[Date of final disposal for application]	
[Patent number]	3407840
[Date of registration]	14.03.2003
[Number of appeal against examiner's decision of rejection]	
[Date of requesting appeal against examiner's decision of rejection]	
[Date of extinction of right]	

THIS PAGE BLANK (USPTO)

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-219835

(43)公開日 平成9年(1997)8月19日

(51)Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 N 5/765			H 0 4 N 5/781	5 1 0 F
	5/781		G 0 6 F 15/40	3 7 0 D
G 0 6 F 17/30			15/401	3 2 0 A
H 0 4 N 5/91			H 0 4 N 5/91	N

審査請求 未請求 請求項の数 4 O L (全 6 頁)

(21)出願番号 特願平8-25343

(22)出願日 平成8年(1996)2月13日

(71)出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72)発明者 谷口 行信

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72)発明者 浜田 洋

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

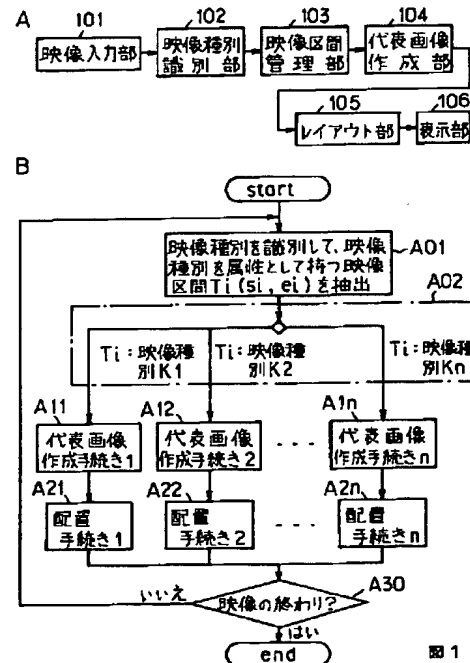
(74)代理人 弁理士 草野 卓

(54)【発明の名称】 映像要約方法および装置

(57)【要約】

【課題】 各映像種別ごとに適した方法で代表画像の抽出を行う。

【解決手段】 映像入力部101より入力される映像データは映像種別識別部102で映像種別が識別され、映像種別を属性として持つ映像区間 $T_i (s_i, e_i)$ が抽出される。抽出された映像区間は映像区間管理部103で管理される。代表画像作成部104では映像種別ごとに予め与えられた手続きが実行され、各映像区間 T_i を代表する画像が抽出あるいは合成される。レイアウト部105で代表画像を空間的に並べた映像要約が作成され、表示部106に表示される。映像種別の一つとしてコマーシャル(CM)を選び、CMを属性として持つ映像区間(CMクリップ)に対して最後の画像を代表画像として抽出することもできる。



【特許請求の範囲】

【請求項1】 映像データを解析して映像種別を識別し、その映像種別を属性として持つ映像区間に映像を時間的に分割し、
該映像種別ごとに予め与えられた手続きを実行して映像区間を代表する画像を抽出あるいは合成し、
その代表画像を空間的に並べて映像要約を作成するたとを特徴とする映像要約方法。

【請求項2】 請求項1の映像要約方法において、
該映像種別としてコマーシャル（CM）を識別し、
コマーシャルの属性を持つ映像区間（CMクリップ）に対して最後の画像を代表画像として抽出することを特徴とする映像要約方法。

【請求項3】 請求項2の映像要約方法において、
コマーシャルを識別する手続きが、映像の時間的な変化量を基にして検出されるカット点列の時間間隔を測定する手続きを含むことを特徴とする映像要約方法。

【請求項4】 映像データを解析して映像種別を識別する映像種別識別部と、
映像種別を属性として持つ映像区間を管理する映像区間管理部と、
該映像種別ごとに予め与えられた手続きを実行することで映像区間を代表する画像を抽出あるいは合成する代表画像作成部と、
代表画像を空間的に並べて映像要約を作成するレイアウト部と、
該映像要約を表示する表示部と、
を具備することを特徴とする映像要約装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は映像データベース、ビデオデッキ、映像編集装置等の映像関連装置において、映像の検索あるいは内容概略把握の効率化を図るための映像要約方法および装置に関する。

【0002】

【従来の技術および発明が解決しようとする課題】ビデオデッキを使ってテープの中から所望の画面を探し出すためには、早送り巻き戻し操作を繰り返すしかなく、時間がかかるという問題点があった。映像データベースなどで大量の映像データを蓄積しておき、それを効率的に利用できるようにするためには、映像を要約してユーザーに提示する機能が必要である。映像の時間的な輝度あるいは色分布の変化に基づいて場面の切り替え（カット）を検出し、場面ごとに1枚あるいは複数枚の画像を抽出して紙の上に一覧表示する方法が開示されている

（外村、谷口、阿久津：Paper Video：紙を用いた新しい映像インタフェース、信学技報、LE94-59、pp.15-20、1994.）。この方法には、映像が長くなると場面の数が多くなるために必ずしも効率的な要約となっていないという問題があった。加えてニュース、コマーシャル

（CM）、映画、天気予報と言った映像種別によって本来、代表画面を抽出する基準が異なるはずであるのに従来方法ではその点を考慮していない。

【0003】映像種別を識別し、映像区間を分割する方法については文献（Doborah Swanberg, Chiao-Fe Shu, Ramesh Jain: Knowledge Guided Parsing in Video Data bases, SPIE Vol.1908 (1993) pp.13-24）に開示されている。この方法は、ニュース映像を対象とし、ニュース映像に特有の空間的知識（ニュースキャスタや文字領域の位置関係）と時間的知識（オープニング、ニューストビックス、天気予報の時間的前後関係）を予め与えておき、画像処理の結果得られる特徴量と与えられた知識とを照合するものである。この方法により、ニュース映像をニューストビックスや天気予報といった映像種別を属性として持つ映像区間に分割することができる。しかし、この文献には要約作成の方法については述べられていない。

【0004】本発明の目的は、上記問題点を解決し、映像種別に応じた方法で代表画像の抽出を行い、より効率的な映像予約を行う方法および装置を提供することにある。

【0005】

【課題を解決するための手段】本発明では、映像データを解析して映像種別を識別し、映像種別を属性として持つ映像区間に映像を時間的に分割し、映像種別ごとに予め与えられた手続きを実行し、映像区間を代表する画像を抽出あるいは合成し、その代表画像を空間的に並べて映像要約を作成する。

【0006】

【作用】請求項1の映像要約方法では、映像データを解析して映像種別を識別し、映像種別を属性として持つ映像区間に分割し、映像種別ごとに予め与えられた手続きに従って代表画像を抽出あるいは合成するので、映像種別に適した方法で代表画像を作成でき、より効率的な要約が作成できるようになる。

【0007】請求項2の映像要約方法では、コマーシャルを映像種別の一つとして識別し、コマーシャルについてはその末尾の画像を代表画像とする。コマーシャルにおいては、その末尾に会社名、商品名などの情報が表示されることが一般的であるので、コマーシャルの末尾の画像を代表画像として抽出することが好適である。

【0008】

【発明の実施の形態】図1の実施例を参照して発明の実施の形態を説明する。図において、映像データ入力部101は映像データを映像種別識別部102に送る。映像データ入力部101はアナログ映像信号をデジタル化する装置であったり、圧縮符号化されたデジタルデータを復号する映像データ復号装置であったりする。映像データには映像データ、音声データおよび撮影時刻に関するタイムコード等の付属データが含まれる。

【0009】映像種別識別部102は101から送られてくる映像データから映像種別を識別して、連続して同じ映像種別を持つ区間を映像種別とともに映像区間管理部103に送る。映像種別としては様々なものが考えられるが、テレビ映像が対象であればコマーシャル、番組、番組の中には映画、ニュース等がある。付属情報の一つとして映像種別が含まれる場合にはそれを利用して映像種別を識別し、映像区間とその種別を出力すればよい。また、付属情報として映像種別が得られない場合には、映像処理によって映像種別を識別する方法を使うこともできる。CM区間の識別方法の一例を後述する。

【0010】103の映像区間管理部では映像種別を属性として持つ映像区間を管理する。図2は、映像区間管理部での映像区間管理方法の一例を示したものである。映像区間の間には包含関係が存在する場合があるので、その包含関係を図2に示すような階層（レイヤー）構造で管理する。例えば、レイヤー203のショット（一つのカメラで撮影された一つの画面）とレイヤー202のCMクリップ（1本のCM）の間には包含関係がある。つまり、CMクリップは複数のショットから構成される。また、放送ではいくつかのCMクリップが連続して現れることが多いのが、この一連のCMをCM区間と呼び、さらに上位のレイヤー201で管理する。レイヤー201でCM区間以外の部分が番組区間である。また、映像中に音楽を含む映像区間、人の声が含まれている映像区間を抽出することもできる。また、各ショットは一連のフレーム205に対応する。

【0011】代表画像作成部104では、映像区間管理部103から映像区間を順に入力し、その映像種別に応じて予め与えられた手続きで代表画像を作成する。例えば、CMクリップという映像種別を属性として持つ映像区間については、映像区間の末尾の映像を代表画像とする。CMクリップにおいては、その末尾に会社名、商品名などの情報が表示されることが一般的であるので、CMの末尾の画像を代表画像として抽出するのが好適である。もちろん、末尾から一定時間前の画像をとるようにすることも、末尾からもっと近い字幕を含むフレームをとるようにすることもできる。番組という属性を持つ映像区間については、映像区間の先頭画像を代表画像とする。番組では最初にタイトルが表示されることが多いので、映像区間の先頭画像を代表画像とするのが望ましい。あるいは番組種別を表す記号を代表画像としたり、代表画像を作成しないようにすることもできる。

【0012】レイアウト部105では、104で作成された代表画像を表示部106の制限（画面の大きさなど）に応じて配置する。例えば、図5に示すように時間順に並べて出来上がったデータを表示部106に送る。表示部106は、コンピュータの画面やプリンタなど画像を表示することのできるデバイスであれば何でもよい。図5に表示の一例を示す。ここでは、CMが一枚の

代表画像503で、番組区間が記号504に表示されており、この画面を見ることでどのようなCMが映像に含まれているかが一目で把握できる。またこの代表画像の上をマウス等で指示すると、その指示がビデオ再生手段に送られ映像の再生を開始するようにすることもできる。またこの画面には501、502に示すようなボタンを使ってユーザが代表画像の作成方法を選択できるようにすることもできる。

【0013】図1Bは図1Aの映像要約装置の処理フローチャートである。映像データを時間順に解析して映像種別を識別し、映像の先頭から数えてi番目に現れる映像区間を T_i とする。映像種別としてはコマーシャル、番組、番組の中にはニュース、映画、天気予報等様々なものが考えられるが、ここでは映像種別として K_1, K_2, \dots, K_n のn種類が識別されるものとする。映像種別 K_j ($j=1 \sim n$)のいずれかを属性として持つ映像区間（同じ種別 K_j の連続する区間） T_i ($i=0, 1, 2, \dots$)に映像を時間的に分割し、 T_i 区間の始点、終点の時刻をそれぞれ s_i, e_i とする（ステップA01）。例えば、テレビ放送では番組とCMが交互に現れ、 T_0 :ニュース、 T_1 :CM、 T_2 :CM、 T_3 :CM、 T_4 :映画、…のようになる。ステップA02では、各映像区間 T_i はその映像種別 K_j に応じて処理が分岐する。 T_i が映像種別 K_1 であれば、ステップA11の代表画像作成手続き1の処理に分岐する。代表画像作成手続き1の処理では、映像区間(s_i, e_i)を代表する画像を抽出あるいは合成する。この代表画像は1枚とは限らず、複数枚であってもよい。続いて、配置手続き1（ステップA21）に処理が移る。配置手続き1では代表画像をどのように空間的に配置するかを決定する。ステップA30では映像の終わりか否かを判定し、終わりでなければステップA01に分岐する。

【0014】請求項2の発明では、ステップA01が映像種別 K_j の一つとしてコマーシャル（CM）を識別し、例えばステップA11がコマーシャルの属性を持つ映像区間（CMクリップ）に対して最後の画像、すなわち時刻 e_i の画像とすることを特徴とする。もちろん、時刻 e_i の代わりに e_i から一定時間さかのぼった時刻の画像を抽出することもできる。

【0015】次に、CM区間を識別する方法について説明する。CM区間が次のような性質を持つことを利用する。

(1) 1本のCM（CMクリップ）は15秒あるいは30秒の長さを持つ（すなわち、CMクリップの開始時刻、終了時刻の差が15秒あるいは30秒である。）

(2) CMクリップ中にはカットが多発することが多い。

【0016】

(3) CM区間は1分程度以上連続してあらわれる。

(4) CMクリップとCMクリップの境界には無音区間

10

20

30

40

50

がある。

まずCMクリップとCMクリップの切れ目の候補を見つけるために、まず、映像の中で画像の内容が大きく変化するところ(カット)を検出する。例えば「大辻、外村、大庭：輝度情報を使った動画像ブラウジング、電気情報通信学会技術報告、IE-90-103,1991」に示されている方法を用いて検出する。検出されたカット時刻の列を{C1, C2, C3, ...}とし、無音区間系列をS1, S2, S3, ...とする。無音区間系列の要素Stは無音区間の開始時刻、終了時刻を属性として持つ。カット

系列Cと無音区間系列SからCM区間を推定する手続きを図3に示す。
【0017】 $t=1, 2, 3$, について、次の処理を行う。まず、カット時刻Ctが無音区間Sに含まれているかどうか調べる(ステップ302)。上述した性質(4)からCMクリップとCMクリップの境界には無音区間があるので、 $Ct \in S$ でないならばCtはCMクリップの先頭でないと判断する。さらに、 $t' > t$ かつ $Ct' - Ct = 15$ (秒)あるいは30(秒)、かつ $Ct' \in S$ を満たす t' が存在するか調べ、存在しなければCtはCMクリップの先頭ではないと判断する(ステップ303)。これは性質(1), (4)を満足していることになる。さらに $t' - t \geq 3$ を満たすか調べ、満たさなければやはりCtはCMクリップの先頭ではないと判断する(ステップ304)。これは性質(2)を満足するか調べていることになる。区間[Ct, Ct']をCM候補区間としてキューに挿入する(ステップ305)。キューの中に60秒以上継続するCM候補区間(すなわち、CM候補区間が60秒以上切れ目なくつながっているもの)が存在するかどうかが調べ(ステップ306)、存在すればその区間をCM区間として出力する(ステップ307)。ステップ306は性質(3)を満足するかどうかが調べていることになる。
【0018】図4にCM区間推定の模式図を示す。401は時間軸に並べられた無音区間系列Sを示し、402*

*はカット系列Ctを示す。カット系列の中で、カットとカットの間の時間間隔が15秒あるいは30秒であり、かつ両端のカットがともに無音区間に含まれるものを403のCM候補区間として抽出する。さらにCM候補区間が60秒以上継続しているものを404のCM区間として出力するわけである。405のCM候補区間は継続時間が60秒未満であったので、CM区間としては出力されない。カット時刻Ctは誤差を含んでいるので303の時間測定ではその誤差を見込んで幅を持った判定を行う方がよい。さらにカットの誤検出、検出もれを見込んでステップ306の判定では60秒以上継続していなくても、その間でカット頻度が高ければCM区間であると判定するようにしてもよい。

【0019】

【発明の効果】以上説明したように、本発明によれば映像種別ごとに代表画像作成の手続きを用意しておくことにより、映像種別に適した方法で代表画像を作成できるようになり、結果的に効率的な要約を作成できるようになる効果がある。また、請求項2および3の発明では、映像種別としてコマーシャルと番組を識別し、コマーシャルの映像区間に対してコマーシャルの最後の画像を代表画像として抽出することによって、コマーシャルの要約リストが自動的に作成できる効果がある。

【図面の簡単な説明】

【図1】Aはこの発明の映像要約装置の実施例を示すブロック図、BはAの処理フローチャート。

【図2】図1Aの映像区間管理部103のデータ管理方法を説明するためのタイミングチャート。

【図3】図1Aの映像種別識別部102におけるCM区間検出処理のフロー図。

【図4】CM区間検出処理を説明するためのタイミングチャート。

【図5】図1の表示部106に表示されたCMを一覧表示した要約表示画面の図。

【図4】

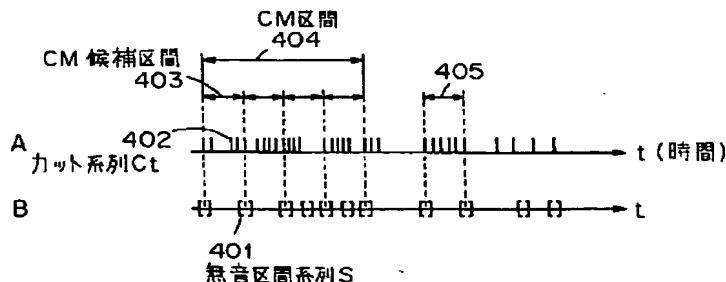


図4

【図1】

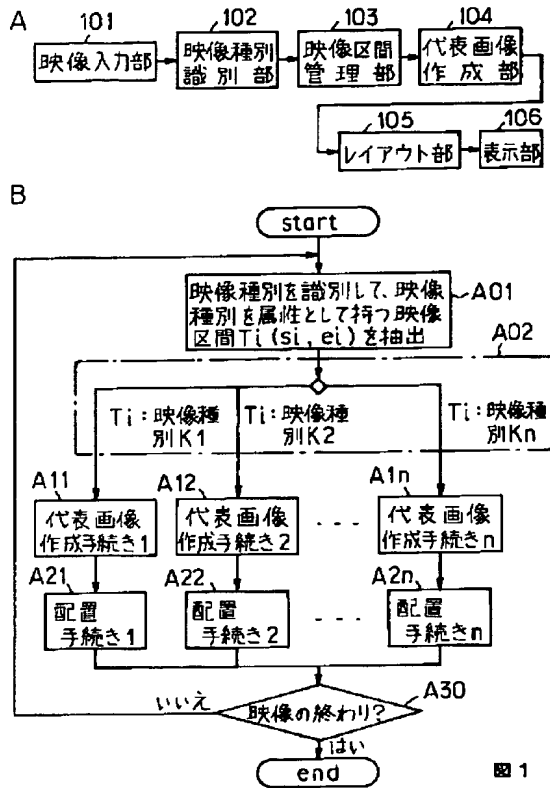


図1

【図3】

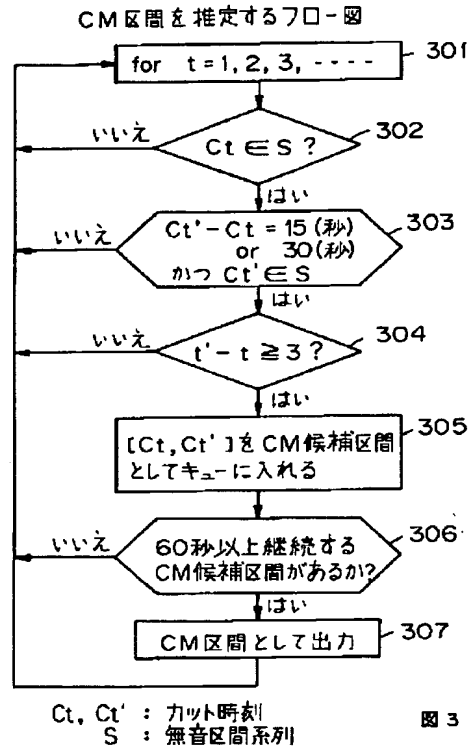


図3

【図2】

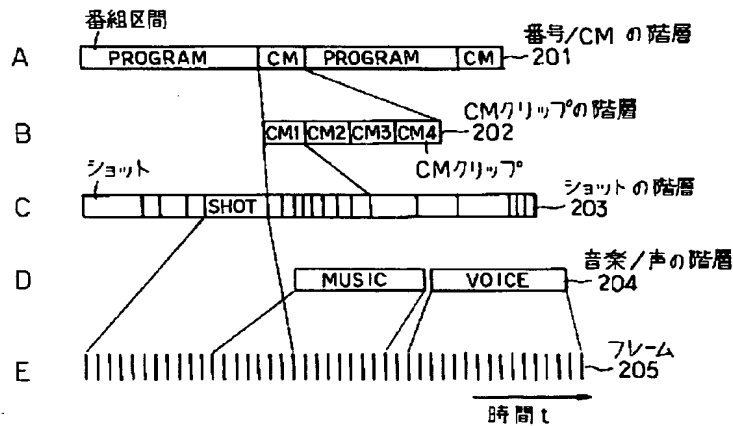


図2

【図5】

